2024 20th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob) | 979-8-3503-8744-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/WIMOB61911.2024.10770523

Integrated Sensing and Communication for Efficient Edge Computing

Khandaker Foysal Haque†, Francesca Meneghello‡ and Francesco Restuccia†
† Institute for the Wireless Internet of Things, Northeastern University, United States
‡ Department of Information Engineering, University of Padova, Italy
E-mails: {haque.k, f.restuccia}@northeastern.edu, francesca.meneghello.1@unipd.it

***Abstract*—Emerging mobile virtual reality (VR) systems are required to continuously perform complex computer vision tasks needing computational power that is excessive for mobile devices. Thus, techniques based on wireless edge computing (WEC) have been recently proposed. However, existing WEC methods require the transmission and processing of a high amount of video data which may ultimately saturate the wireless link. In this paper, we propose a novel *sensing-assisted edge computing* (`ISAC-EC`) approach to address this issue. `ISAC-EC` leverages knowledge about the physical environment to reduce the end-to-end latency and overall computational burden by transmitting to the edge server only the relevant data for the delivery of the service. Our intuition is that the transmission of the portion of the video frames where there are no changes with respect to the previous frames can be avoided. Through wireless sensing, only the part of the frames where any environmental change is detected is transmitted and processed. We evaluated `ISAC-EC` by using a 10K 360°camera with a Wi-Fi 6 sensing system operating at 160 MHz and performing localization and tracking. Experimental results show that `ISAC-EC` reduces both the channel occupation and end-to-end latency by more than 90% while improving the instance segmentation and object detection performance with respect to state-of-the-art WEC approaches. For reproducibility purposes, we pledge to share our dataset and code repository.**

I. Introduction

Emerging technologies based on mobile virtual reality (VR), such as the Metaverse, will provide new entertainment applications, and ultra-realistic online learning experiences among others. One of the key issues currently stymieing the Metaverse is that commercial VR headsets do not deliver adequate performance to the end user. Experts believe that 360° video frames should have at least 120 Hz frame rate with 8K resolution to avoid pixelation and motion sickness [1]. However, current wireless VR headsets on the market provide up to 4K resolution, with only a limited few achieving a frame rate within the range of 100-120 Hz [1]. Moreover, Existing mobile VR headsets do not have enough computational resources to execute the required complex deep neural network (DNN) tasks like object detection and segmentation on 8K frames. Thus, they either excessively decrease battery lifetime or degrade the performance to unacceptable levels. While mobile DNNs such as MobileNet and MnasNet can decrease the computation requirements, they lose in accuracy – up to 6.4% – compared to large DNNs such as ResNet-152. While wireless edge computing (WEC) can address the issue, continuously offloading DNN tasks requires data rates that far exceed what existing wireless technologies can offer today. Indeed, sending frames at 120 Hz with 8K resolution would require about 40 Gbps of data rate for each AR/VR device, while today, Wi-Fi supports a maximum of 1.2 Gbps network-wide [2].

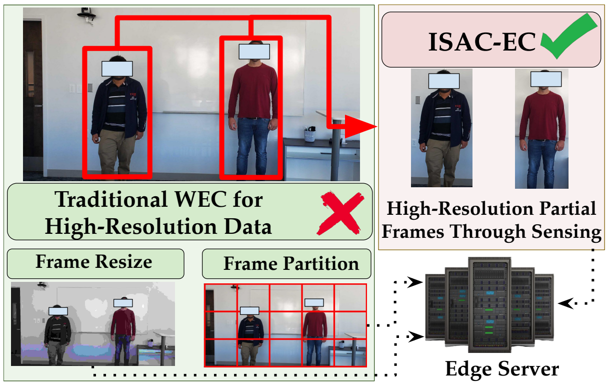


Fig. 1: `ISAC-EC` vs traditional edge computing approaches.

Another key issue is that existing DNNs are not trained on 8K images due to the unavailability of the dataset and limited computing resources. While compressing and/or downsizing the frames would reduce the data rate requirement, they would reduce the system performance by 79% and 62% respectively [2]. While partitioning the 8K frames into multiple smaller tiles can be another option, the region of interest might fall into multiple tiles causing performance degradation.

To address the above core issues, we propose *Integrated Sensing Assisted Efficient Edge Computing* (`ISAC-EC`). As depicted in Figure 1, `ISAC-EC` offloads to the edge server only the relevant portions of the frames – thus decreasing data rate requirements with respect to traditional edge computing approaches, without losing any image resolution. In stark opposition with existing art that compresses or partitions frames [3], `ISAC-EC` leverages wireless sensing to localize and track environmental changes such as the movements of humans or other objects. This way, instead of offloading the whole frame, `ISAC-EC` only offloads the part of the frame where motion is detected, hereafter referred to as region of interest (ROI). Since wireless sensing-based localization can operate concurrently with wireless communications, `ISAC-EC` does not require any dedicated infrastructure as it leverages the channel estimation procedure which is routinely required by any wireless communication standard such as Wi-Fi (IEEE 802.11). Even though it would be possible to take similar approaches with event-based cameras or software-only-vision [4], the approach in `ISAC-EC` is much simpler which also does not need any dedicated additional device.

Summary of Novel Contributions

• We present `ISAC-EC`, a novel paradigm that performs wireless edge computing by leveraging wireless sensing integrated with the communication process (Section III). `ISAC-EC` optimizes the edge offloading process by localizing and tracking environmental changes that eventually determine the ROI to offload to the edge server instead of the whole frame. This way, ISAC-EC optimizes the transmission latency and channel utilization while improving the DNN performance;

- We implemented and validated the ISAC-EC through several experiments carried out in a hall room (Section V). In Section VI, we compare ISAC-EC with state-of-the-art work YolactACOS [5] and EdgeDuet [3].

II. Background and Related Work

Wireless edge computing (WEC) has gained significant traction over the last few years. Prior work has focused on reducing the end-to-end latency as well as the channel occupation by employing DNN partitioning frame partitioning and full task offloading [6]. While frame downsizing and frame compression approaches [6] can effectively decrease energy consumption and channel usage, they can hardly be applied in VR applications with 8K resolution and higher. Some existing work that is complementary to ours has used wireless sensing for vision-oriented approaches in multi-modal, cross-modal, and transfer learning settings [7]. For example, Xie et al. [8] rather than leveraging angle of arrival (AoA), utilizes a single off-the-shelf time of arrival (ToA) based depth camera to generate high-resolution maps in a noisy and dark environment.*However, to the best of our knowledge, none of the earlier work has proposed integrated wireless sensing to assist efficient edge offloading.*

III. System Overview

ISAC-EC empowers modern mobile devices featuring 360° cameras with wireless sensing functionalities that identify the ROI inside each captured video frame before transmitting it to the edge. ISAC-EC consists of three main blocks, as summarized in Figure 2: (1) sensing-assisted ROI detection, (2) ROI offloading, and (3) task execution at the edge server.

(1) Sensing-assisted ROI Detection: The ROI detection is based on the context information obtained through wireless sensing. Specifically, we leverage the channel frequency response (CFR) estimated by the wireless network interface card (NIC) to detect the dynamics in the environment. ISAC-EC synchronizes channel measurements with the video frames through their timestamps, and obtains an estimate of the locations of the targets in the environment by processing the CFR through multi-path parameter estimation algorithms. A tracking algorithm allows detecting the ROI, being the area where changes were detected.

(2) ROI Offloading: The ROI offloading block is in charge of selecting the portion of the high-resolution frame to be offloaded to the edge server based on the context information gained through wireless sensing.

(3) Task Execution at Edge Server: The edge server receives the ROIs from the camera and uses them as input for the DNN task, thus reducing both the training and inference time with respect to using the whole (bigger) frames.

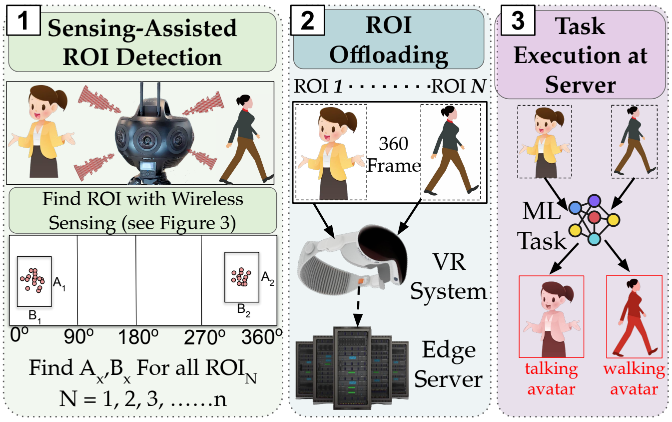


Fig. 2: Overview of the ISAC-EC Framework.

IV. Sensing-Aided ROI Detection

The CFR-based localization has been implemented by adapting the super-resolution multi-path parameter estimation algorithm MD-Track [9]. Hence, for tracking, we implemented a custom-tailored approach based on density-based spatial clustering of applications with noise (DBSCAN).

A. AoA and ToA Estimation

We consider a $1 \times N$ system where N is the number of receiving antennas and $n \in \{0, \ldots, N-1\}$ indicates the receiving antenna index. Being f_c the main carrier frequency, Δf the orthogonal frequency-division multiplexing (OFDM) sub-channel spacing and $T = 1/\Delta f$ the OFDM symbol time, the CFR for sub-channel $k \in \{-K/2, \ldots, K/2-1\}$, estimated at receiver antenna n and time t, $H_{k,n}(t)$, is modeled as

$$H_{k,n}(t) = \sum_{p=0}^{P-1} A_p(t) e^{-j2\pi(f_c + k/T)\tau_{p,n}(t)} , \tag{1}$$

where $p \in \{0, \ldots, P-1\}$ represents the P multi-path components associated with the wireless signal propagation, each of which is characterized by an attenuation $A_p(t)$ and a ToA $\tau_{p,n}(t)$ (also referred to as propagation delay). Each multi-path component p is associated with a static or moving object in the environment that acts as a reflector, diffractor, or scatterer for the wireless signal propagating from the transmitter to the receiver. The propagation delay $\tau_{p,n}(t)$ is associated with the position $\ell_p(t)$ of the p-th object in the environment and the collecting antenna. Each multi-path component is collected by each antenna in subsequent time instants that depend on the AoA $\theta_{\mathrm{rx},p}(t)$. Indicating with $\Delta^{\mathrm{rx}}_{p,n}(t)$ the antenna-dependent contribution to the length of the p-th component, the ToA is obtained as

$$\tau_{p,n}(t) = \frac{\ell_p(t) + \Delta^{\mathrm{rx}}_{p,n}(t)}{c} , \tag{2}$$

where c is the speed of light. Considering a linear array with antennas spaced apart by d_{rx} and using the left antenna as the reference, $\Delta^{\mathrm{rx}}_{p,n}(t) = n \sin(\theta_{\mathrm{rx},p}(t)) d_{\mathrm{rx}}$, where the AoA $\theta_{\mathrm{rx},p}(t)$ is measured clockwise starting from the direction perpendicular to the antenna array. In this work, we use the iterative mD-Track algorithm for this purpose [9]. Note that we focus on the

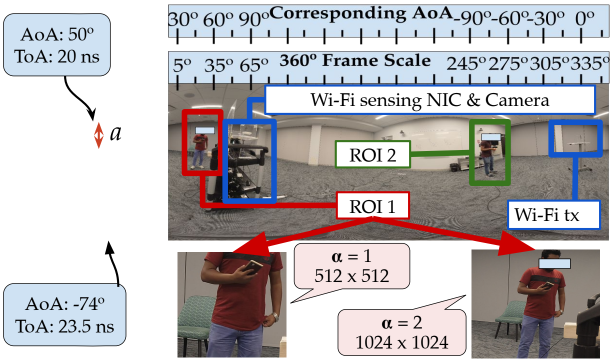


Fig. 3: ROI detection from AoA/ToA estimation and clustering. The ROI multiplying factor α is the parameter used for obtaining the size of the ROI from the extension of the cluster (a) to account for localization errors.

identification of the AoA and ToA (2-dimensional mD-Track) as they are sufficient for providing enough side information for ROI detection. After that, the AoA/ToA pairs are clustered through density-based spatial clustering of applications with noise (DBSCAN), removing the outliers associated with noise. Hence, the centroids are compared with the centroids of the clusters associated with the previous video frame to detect any change in their location. This information about moving targets is then used to obtain the ROI.

B. AoA and ToA Projection into the Camera Reference System

For proper sensing-assisted ROI selection, the NIC for CFR data collection and the 360° camera have to share the same reference system. This would require: (i) the camera and the sensing NIC to be exactly co-located and, in turn, capture the same field of view, and (ii) the NIC transmitting opportunistic signals to trigger CFR estimation to be placed exactly on the tangent of sensing NIC antenna array, to share the same common zero of AoA reference frame. However, these requirements are hardly achievable from a physical perspective due to the physical location of the devices. In turn, we need to 'project' the AoA and ToA estimated through the sensing NIC into the camera reference system. For this, we designed a procedure consisting of two steps. First, we set a AoA scale for the 360° frame assigning 0° to the left edge of the camera field of view. Hence, we project the AoA $\angle OBD$, obtained above, into the AoA value in the new reference system, hereafter referred to as θ. Specifically, we have $\theta = [(\angle OBD + \theta_{\text{tx}}) \mod 360]$ where mod represents the modulo operation and θ_{tx} indicates the location of Wi-Fi transmitter in the 360° reference system. An example of the processing is depicted in Figure 3 using real experimental data. For example, $\angle OBD = 50°$ identifying ROI 1 corresponds to $\theta = 25°$.

V. Experimental Setup

We evaluated `ISAC-EC` by implementing the system on commercial devices available on the market. The system comprises an Insta360 Titan 360° camera for video capturing, a Linux machine for video processing, and an IEEE 802.11ax network for communication and wireless sensing (localization and tracking). The system was configured to capture frames at a resolution of 10K with a frame rate of 25 frames per second (FPS). The IEEE 802.11ax network comprised two commercial-off-the-shelf (COTS) AX200 NIC operating on a 5 GHz Wi-Fi channel with 160MHz of bandwidth. We placed the Wi-Fi receiver and the 360° camera in closed proximity and synchronized the Wi-Fi localization and camera systems by using their internal reference clock. The Wi-Fi transmitter was placed on the opposite side to properly irradiate the environment and obtain valuable information for sensing. We carried out the experiments deploying the experimental setup in an entrance hall that allows evaluating the performance of `ISAC-EC` in a real-world environment. The experimental setup along with the evaluation scenario is presented in Figure 4. We performed six different tests: a single person walking between (i) H1 to H2 (ii) H3 and H4 (iii) V1 and V2 (iv) V3 and V4, and two persons walking simultaneously between (v) H1 & H2 and H3 & H4 and (vi) V1 & V2 and V2 & V3.

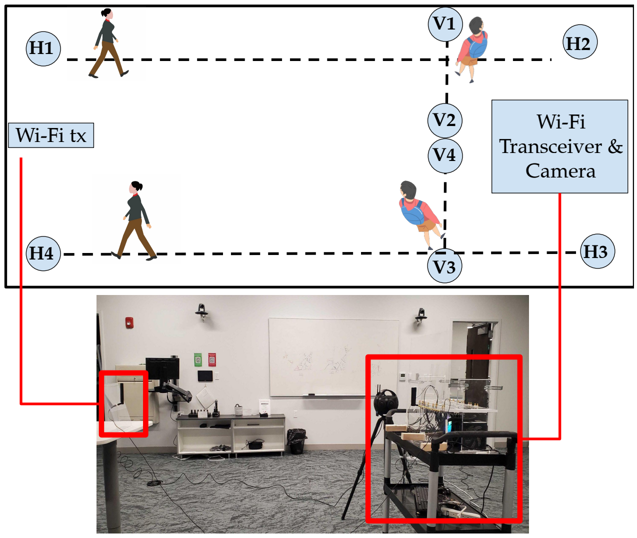


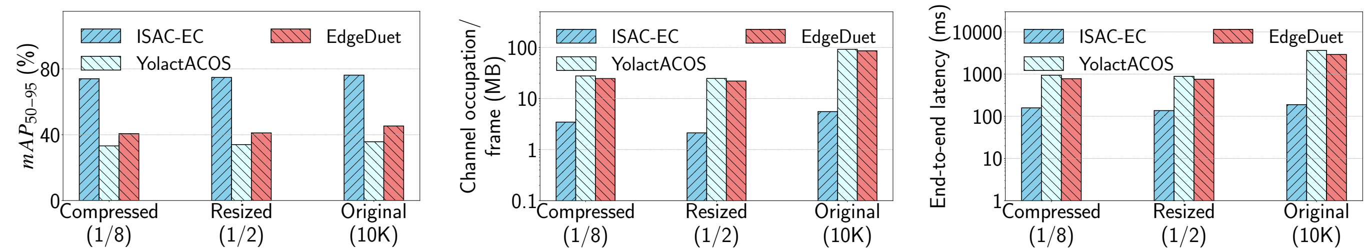
Fig. 4: Experimental setup and `ISAC-EC` evaluation scenario.

VI. Performance Evaluation

We assessed the performance of `ISAC-EC` by performing an extensive experimental data collection campaign and evaluating three main metrics: (i) accuracy of the DNN task; (ii) wireless channel occupation; and (iii) end-to-end latency. In the following results, the metrics have been averaged over all the frames captured. For a baseline comparison, we evaluate `ISAC-EC` against two state-of-the-art (SOTA) algorithms: (i) YolactACOS adaptive edge assisted segmentation [5], (ii) EdgeDuet context-aware data partition-based WEC [3]. For the DNN task we chose *instance segmentation* with YOLOv8m as it is one of the widely used benchmarking tasks for edge computing.

A. Performance for Different Edge Computing Approaches

We also perform the comparative performance analysis of `ISAC-EC`, YolactACOS, and EdgeDuet when the frames for all the approaches are downsized and compressed to $1/2$ and $1/8$ respectively. Figure 5a presents the mAP_{50-95} of the three



(a) mAP_{50-95} of original, resized (1/2), and compressed (1/8) frames.

(b) Channel occupation of original, resized (1/2), and compressed (1/8) frames.

(c) End-to-end latency of original, resized (1/2), and compressed (1/8) frames.

Fig. 5: Comparative analysis of `ISAC-EC` with YolactACOS and EdgeDuet.

approaches with original (10K), resized (1/2), and compressed (1/8) versions of the video frames. The performances for all the approaches degrade slightly due to the rational image resizing and compression. The results show that `ISAC-EC` achieves an mAP_{50-95} of 65% whereas the state-of-the-art approaches– YolactACOS, and EdgeDuet degrade the performance by 41% and 34% on average with 10K image resolutions.

B. Analysis of Channel Occupation Usage

Analyzing the per-frame channel occupation for all the methods – `ISAC-EC`, YolactACOS, and EdgeDuet – is paramount, considering that the radio spectrum is a pivotal and limited resource for wireless networks. The average per-frame channel occupation of `ISAC-EC`, YolactACOS, and EdgeDuet with three different frame sizes – original (no resize and no compression), resized (1/2), and compressed (1/8) – is depicted in Figure 5b. The results show that for original frames, `ISAC-EC` reduces the channel occupation by 94.03%, and 93.59% respectively in comparison to YolactACOS and EdgeDuet. To elaborate further, on average, the size of an original frame is 138.88 MB whereas the average channel occupation required by `ISAC-EC`, YolactACOS, and EdgeDuet are 5.5 MB, 92.2 MB and 85.9 MB per frame respectively.

C. End-to-End Latency Analysis

End-to-end latency is one of the critical factors for time-critical edge computing tasks including a wide range of VR applications. We analyze the end-to-end latency of `ISAC-EC` by comparing it with YolactACOS, and EdgeDuet for different frame types as presented in Figure 5c. The average end-to-end latency of `ISAC-EC` is 188.62 ms, 136.67 ms, and 158.46 ms for original, resized (1/2), and compressed (1/2) frames respectively, which is much lower than the other two approaches. With 10K resolution, `ISAC-EC` improves the latency by 94.80% and 93.52% in comparison to the YolactACOS and EdgeDuet respectively.

VII. Conclusions

In this paper, we proposed a new paradigm for wireless edge computing called `ISAC-EC`. Our new approach leverages Wi-Fi-based localization and tracking to support high resource-consuming 360° computer vision tasks by obtaining the location of ROI based on the environment dynamics. This information allows offloading to the edge server only the detected ROI instead of the entire frame, thus reducing airtime overhead and overall latency. Our proposed approach reduces the overall end-to-end latency by 94.80% and 93.52% respectively while achieving 39.69% and 23.64% net mAP_{50-95} improvement in comparison to the SOTA WEC approach – YolactACOS, and EdgeDuet – in image segmentation tasks.

Acknowledgements

This work has been funded in part by the National Science Foundation under grants CNS-2134973, CNS-2120447, and ECCS-2229472, by the Air Force Office of Scientific Research under contract number FA9550-23-1-0261 and by the Office of Naval Research under award number N00014-23-1-2221. The work was also partially supported by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership on "Telecommunications of the Future" (PE00000001 - program "RESTART"), and by the Fulbright Schuman Program, administered by the Fulbright Commission in Brussels and jointly financed by the U.S. Department of State, and the European Commission.

References

[1] A. Somrak, I. Humar, M. S. Hossain, M. F. Alhamid, M. A. Hossain, and J. Guna, "Estimating VR Sickness and User Experience Using Different HMD Technologies: An Evaluation Study," *Future Generation Computer Systems*, vol. 94, pp. 302–316, 2019.

[2] K. F. Haque, F. Meneghello, M. E. Karim, and F. Restuccia, "Sawec: Sensing-assisted wireless edge computing," *arXiv preprint arXiv:2402.10021*, 2024.

[3] Z. Yang, X. Wang, J. Wu, Y. Zhao, Q. Ma, X. Miao, L. Zhang, and Z. Zhou, "Edgeduet: Tiling Small Object Detection for Edge Assisted Autonomous Mobile Vision," *IEEE/ACM Transactions on Networking*, 2022.

[4] S. D. Uhlrich, A. Falisse, Ł. Kidziński, J. Muccini, M. Ko, A. S. Chaudhari, J. L. Hicks, and S. L. Delp, "OpenCap: Human Movement Dynamics from Smartphone Videos," *PLoS computational biology*, vol. 19, no. 10, p. e1011462, 2023.

[5] Y. Xie, Y. Guo, Z. Mi, Y. Yang, and M. S. Obaidat, "Edge-Assisted Real-Time Instance Segmentation for Resource-Limited IoT Devices," *IEEE Internet of Things Journal*, vol. 10, no. 1, pp. 473–485, 2022.

[6] C. Feng, P. Han, X. Zhang, B. Yang, Y. Liu, and L. Guo, "Computation Offloading in Mobile Edge Computing Networks: A Survey," *Journal of Network and Computer Applications*, vol. 202, p. 103366, 2022.

[7] Y. Zhao, J. Xu, J. Wu, J. Hao, and H. Qian, "Enhancing Camera-based Multimodal Indoor Localization with Device-Free Movement Measurement using WiFi," *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 1024–1038, 2019.

[8] Z. Xie, X. Ouyang, L. Pan, W. Lu, G. Xing, and X. Liu, "Mozart: A Mobile ToF System for Sensing in the Dark through Phase Manipulation," in *Proceedings of the 21st Annual International Conference on Mobile Systems, Applications and Services*, 2023, pp. 163–176.

[9] Y. Xie, J. Xiong, M. Li, and K. Jamieson, "MD-Track: Leveraging Multi-Dimensionality for Passive Indoor Wi-Fi Tracking," in *Proc. of ACM MobiCom*, 2019.